EX PARTE OR LATE FILED **BELLSOUTH**

Kathleen B. Levitz
Vice President-Federal Regulatory

Suite 900
1133-21st Street, N.W.
Washington, D.C. 20036-3351
202 463-4113
Fax: 202 463-4198
Internet: levitz.kathleen@bsc.bls.com

RECEIVED
MAY 20 1999
FEDERAL COMMUNICATIONS COMMISSION
OFFICE OF THE SECRETARY

May 20, 1999

EX PARTE

Ms. Magalie Roman Salas
Secretary
Federal Communications Commission
The Portals
445 12th St. SW
Washington, D.C. 20554

Re: CC Docket No. 98-56 and CC Docket No. 98-121

Dear Ms. Salas:

This is to inform you that on May 19, 1999 Venetta Bridges and I, representing BellSouth, and Dr. Fritz Scheuren and Dr. Edward Mulrow of Ernst & Young met with members of the Common Carrier Bureau staff. The following Common Carrier Bureau staff members attended at least part of the meeting: Alex Belinfante; Whitey Thayer; Florence Setzer; and Daniel Shiman.

During the meeting, BellSouth representatives gave a status report on the workshops that the Louisiana Public Service Commission ("LPSC") staff held on in LPSC Docket No. U22252 – Subdocket C). The purpose of these workshops is to identify the performance measurements, standards and statistical analyses that the LPSC should use to determine whether BellSouth is meeting its statutory obligation to provide CLECs with nondiscriminatory access to UNEs and services. We then focused upon the efforts of Dr. Scheuren, Dr. Mulrow, and his staff at Ernst & Young to respond to written questions posed by the Bureau staff on April 12, 1999. The attached documents formed the basis for that presentation.

Because the Commission has been considering issues related to performance measurements and standards in both proceedings identified above, we are filing notice of this ex parte meeting in both dockets, as required by Section 1.1206(b)(2) of the Commission's rules. Please associate this notice with the record of both dockets.

Sincerely,

Kathleen B. Levitz

Kathleen B. Levitz
Vice President – Federal Regulatory

Attachments

cc: Alex Belinfante (w/o attachment)
Florence Setzer (w/o attachment)
Daniel Shiman (w/o attachment)
Whitey Thayer (w/o attachment)

Open Issues

Comparing Like-to-Like

AGREEMENT: The need for like-to-like comparisons requires the data to be disaggregated to a very deep level. This includes wire center and time of month, as well as SQM disaggregation levels defined by the Louisiana Public Service Commission.

Open Issues

Performance Measure Test Statistic

AGREEMENT: Each performance measure of interest should be summarized by one overall test statistic giving the decision maker a rule that determines whether a statistically significant difference exists.

Open Issues

Test Statistic Methodology

➢ **BellSouth:**

- Overall service process defines parity
- Testing at aggregate level sufficient to discern favoritism
- SQM-level disaggregation reports available to explore data

➢ **LCUG:**

- Construct indicator at lowest level of disaggregation
- Make allowance for random variation, assuming parity holds
- Aggregate statistic should detect consistent violations in any cell

Open Issues

Type I and Type II Errors

AGREEMENT: Both sides agree that Type I and Type II Error probabilities should be balanced.

In the event that the proposed approach to attain this balance is not workable:

- BellSouth:
  - has proposed a feasible altenative
- LCUG:
  - has proposed no alternative solution

EY econ
Economics Consulting & Quantitative Analysis
EY/Econ-STAT

Open Issues

Statistical Paradigm

AGREEMENT: The system must be developed so that it can be put into production mode.

- BellSouth:
  - Methodology is in production mode
- LCUG:
  - Current methodology is not in production mode

econ
Economics Consulting & Quantitative Analysis
EY/Econ-STAT

Paradigms of Analysis

- **Exploratory Paradigm**
  - Develop understanding of measure and underlying data
  - Investigate anomalies in data
  - Tailor methodology specifically to data set

- **Production Paradigm**
  - Use rote analysis methodology
  - Process data
  - Automated production of monthly reports
  - Document process and results

Paradigms of Analysis

➢ Exploratory Paradigm

- ***PROS***
  - Examine underlying data
  - Protects against using erroneous data
- ***CONS***
  - Time consuming
  - Not feasible on continuing basis

➢ Production Paradigm

- ***PROS***
  - Fast, automated reports
  - Documentation of results
- ***CONS***
  - No investigation of data anomalies

EY econ
Economics Consulting & Quantitative Analysis
EY/Econ-STAT

Open Issues

Trimming

AGREEMENT: Trimming is needed but finding a robust rule that can be used in a production setting is difficult.

Open Issues

Independence of Performance Measures

Agreement: Correlation between the performance measures must be accounted for by aggregating over similar measures.

Federal Communications Commission
Washington, D.C. 20554

April 12, 1999

Ms. Kathleen Levitz, BellSouth
Vice-President-Federal Regulatory
BellSouth Corporation
Suite 900
1133-21st Street, N.W.
Washington, DC 20036-3351

Dear Ms. Levitz:

In Appendix B of the Commission's Notice of Proposed Rulemaking regarding performance measurements and reporting requirements for operations support systems ("OSS"), interconnection, and operator and directory assistance (CC Docket No. 98-56), the Commission sought comment on the use of statistical methodologies for evaluating an incumbent local exchange carrier's ("ILEC's") performance in the provisioning of OSS to requesting carriers. A number of parties have proposed various statistical methods that the Commission could use in evaluating an ILEC's provisioning of OSS. As a result of reviewing the various proposals of statistical methodologies, the Bureau Staff has developed a list of questions concerning the proposals made by the Local Competition Users Group ("LCUG") and BellSouth, which are attached. Please provide written responses to these questions at your earliest convenience and file the responses in CC Docket No. 98-56.

If you have any questions, please contact Daniel Shiman at (202) 418-7153.

Sincerely,

Carol E. Mattey, Chief
Policy and Program Planning Division
Common Carrier Bureau
Federal Communications Commission

Questions Concerning the Statistical Methodology to Use for Evaluating Performance Measurements

Please note that this list of questions does not cover all of the issues that have been identified, or that may be considered important.

General Comparison

1. Please compare the BellSouth and LCUG proposed tests (and their proposed variants) according to the following criteria. Provide theoretical analysis (with mathematical proofs), data analysis, examples, and cites to relevant references, where possible.
   - Efficiency (power) of the test to detect discrimination, including higher variance of CLEC data
   - Ability to handle confounding variables
   - Ability to handle heteroscedasticity
   - Ability to handle correlation (dependency) of measures, and correlation of subcells for a measure generated by disaggregating according to confounding variables
   - Ability to handle nonnormality of the data and small sample sizes

Concerning Efficiency:

2. What is the relative power of the BellSouth and LCUG tests? Assuming disaggregation is done according to multiple confounding variables, which will create many disaggregated cells, please give OC (or power) curves (using identical assumptions for both tests) for the following alternative hypotheses (H1-H4):

   Null hypothesis: H0: no discrimination

   vs. Alternative hypotheses:

   H1: discrimination in all cells
   H2: discrimination in half the cells
   H3: discrimination in 25% of the cells
   H4: discrimination in just one cell

   Also examine the two tests' ability to detect discrimination AGAINST the CLEC under the following hypothesis:

   H3a: discrimination against the CLEC in 25% of cells, discrimination for the CLEC in another 25% of cells

   Are there any other alternative hypotheses that should be considered? Is there any reason to believe that one particular scenario of potential discrimination is most likely to occur or most important?

3. How important is it to balance the probability of Type I and Type II errors? Is there a mechanical formula that would adjust the critical values (and hence the probability of a Type I error) as the sample size varied? How can we explicitly measure the costs of a Type I and of a Type II error, as BellSouth suggests needs to be done?

Concerning Estimating the Variance:

4. Why is it desirable to use replication to estimate the variance? What advantages does this have over using an alternative method?

Concerning Aggregating the Data:

5. What are the specific advantages/disadvantages of using aggregation of the adjusted data (the BellSouth approach)? Compare to testing unadjusted aggregate data (LCUG's original proposal) and testing individual cells of disaggregated data (LCUG's recent approach)? In particular, consider the criteria discussed in question 1.

6. Are there tests that can be performed to determine the validity of the degree of aggregation that BellSouth proposes versus the degree of disaggregation LCUG proposes? Is there some middle ground that can be reached through such tests by aggregating some of the cells, where appropriate, and disaggregating where aggregation is not appropriate?

Concerning Dependency:

7. Isn't the replicate estimate of the variance also affected by dependency (i.e., correlation) in the data? This appears to be confirmed in Wolters. Which methodology is affected more by dependency in the data?

8. How much dependency is there in the data (between measures, wirecenters, over time)? How can this be determined? Should this be determined using statistical means, or by examining physical relationships involved between measures (dependence on common computer system or common cable), or by examining each event *ex post*? Can a covariance matrix be developed using weekly or daily data, or by matching the ILEC and CLEC data using the multiple cells created through disaggregation? How much will the dependency measured affect the probability of a Type I error for the LCUG method, and for the BellSouth method?

Concerning Normality of the Data, and Sample Sizes:

9. Are the data nonnormal? How can this be determined? What size sample do we need to get an approximately normal distribution of a mean? How can this be determined?

10. Is permutations testing the best way to handle small, nonnormal samples? What are the advantages and disadvantages of permutations testing? Are there any problems that small sample sizes create for BellSouth's proposed methodology? Could we see a comparison of the results using permutation testing with BellSouth's results?

Concerning Statistical vs. Competitive Significance of the Results:

11. Should a statistically significant difference in means be interpreted to mean that there is discrimination in the process? In other words, should we consider whether the observed difference in means will have an economic impact on CLECs' business? Won't very large sample sizes tend to make even small differences in means statistically significant? How large should a difference in means be for a particular measure for it to be considered "competitively significant" and therefore discriminatory? How should this "threshold difference" be determined for each measure? How can a "threshold difference" be implemented for a testing procedure?

Concerning Confounding Factors:

12. Is it necessary to disaggregate according to every confounding factor? What are the advantages and disadvantages of doing so? Would it be possible to disaggregate only

according to those confounding factors that are statistically determined to have an impact on the results?

General Comparison

- Power to Detect Discrimination
- Confounding Variables
- Heteroscedasticity
- Dependency
  - Correlation of Measures
  - Correlation of Subcells
- Non-normality
- Small Sample Sizes

EY econ
Economics Consulting & Quantitative Analysis
EY/Econ-STAT

Test Hypotheses

- Null Hypothesis:

$$\mu_{ILEC} = \mu_{CLEC}$$
$$\sigma^2_{ILEC} = \sigma^2_{CLEC}$$

- Alternative Hypothesis:

$$\mu_{CLEC} = \mu_{ILEC} + d\sigma_{ILEC} \quad (d > 0)$$
$$\sigma^2_{CLEC} = r\sigma^2_{ILEC} \quad (r \geq 1)$$

We only consider r = 1 in what follows.

Testing Errors

- Type I: Rejecting the null when it is true.
- Type II: Accepting the null when it is false.

Test Statistics

Test	Formula
Modified Z	$\dfrac{\bar{x}_1 - \bar{x}_2}{s_1\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$
Adjusted Modified Z	$\dfrac{\bar{x}_{1w} - \bar{x}_2}{s_{1w2}}$
Adjusted Jackknife	$\dfrac{\hat{\hat{D}}}{\sqrt{v(\hat{\hat{D}})}} \cdot \dfrac{\sqrt{c_1 s_{1w}^2 + \dfrac{s_2^2}{n_2}}}{s_{1w}\sqrt{c_1 + \dfrac{1}{n_2}}}$

See Handout for explanation of terms.

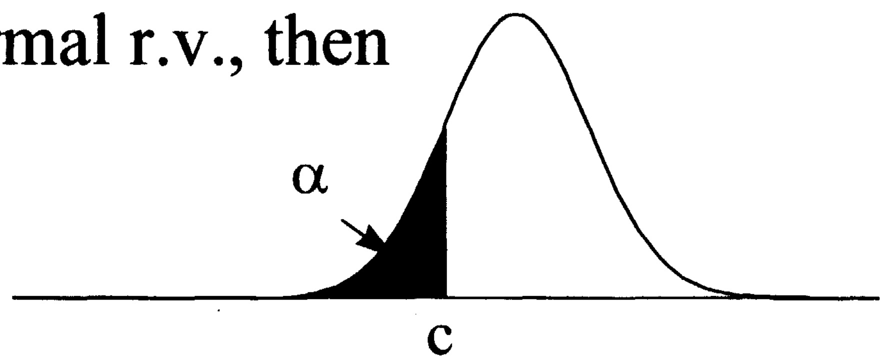
Error Probabilities

- Denote the test statistic by Z and the critical value of the test by c. Given the null hypothesis is true:

  Type I: $\alpha = P(Z < c)$

  α is also called the significance level of the test.

- If, under the null hypothesis, Z is a standard normal r.v., then



Economics Consulting & Quantitative Analysis
EY/Econ-STAT

Error Probabilities

Type II: Given a particular alternative is true $\beta = P(Z \geq c)$

If the observations are independent then

$$Z = \frac{\overline{X}_{ILEC} - \overline{X}_{CLEC} + \sigma_{ILEC}\, d}{\sigma_{ILEC}\sqrt{\frac{1}{n_1} + \frac{r}{n_2}}}$$

$$\beta = P\left(Z \geq \frac{c\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} + d}{\sqrt{\frac{1}{n_1} + \frac{r}{n_2}}} \right)$$

Economics Consulting & Quantitative Analysis
EY/Econ-STAT

Power

Power of a Test: $1 - \beta$

Depends on:

- Critical value (or equivalently the significance level)
- ILEC and CLEC sample sizes
- The alternative hypothesis (values of d and r)

➢ For Jackknife we need to simulate the process to evaluate the power.

EY econ
Economics Consulting & Quantitative Analysis
EY/Econ-STAT

Simulation Procedure

- 240 wire centers with 2 classes each are created.
- Sample sizes are randomly generated and randomly allocated across subclasses. Total sample size has a mean of 29,120 of which 5% are, on the average, CLEC orders.
- All subclass means and standard deviations are generated based on wire center and class effects.

$$\mu_{jk} = \mu + u_j + v_k$$

$$\sigma_{jk} = t_j w_k \sigma$$

j = 1, … 240 wire centers
k = 1,2 classes

EY econ
Economics Consulting & Quantitative Analysis
EY/Econ-STAT

Simulation Procedure

- μ and σ are the overall mean and variance (= 0 and 1 respectively).
- u_j, v_k, t_j and w_k are randomly generated.
- Correlation coefficients are randomly generated for each wire center. Orders within a wire center are generated so that the subclass ILEC - CLEC mean differences are correlated within a wire center.
- This is equivalent to generating a specific type of random effect model.

(see handout for complete details)

EY econ
Economics Consulting & Quantitative Analysis
EY/Econ-STAT

Simulation Results

1000 Monte Carlo Runs

Type I Errors
(Critical value for tests = -1.65)

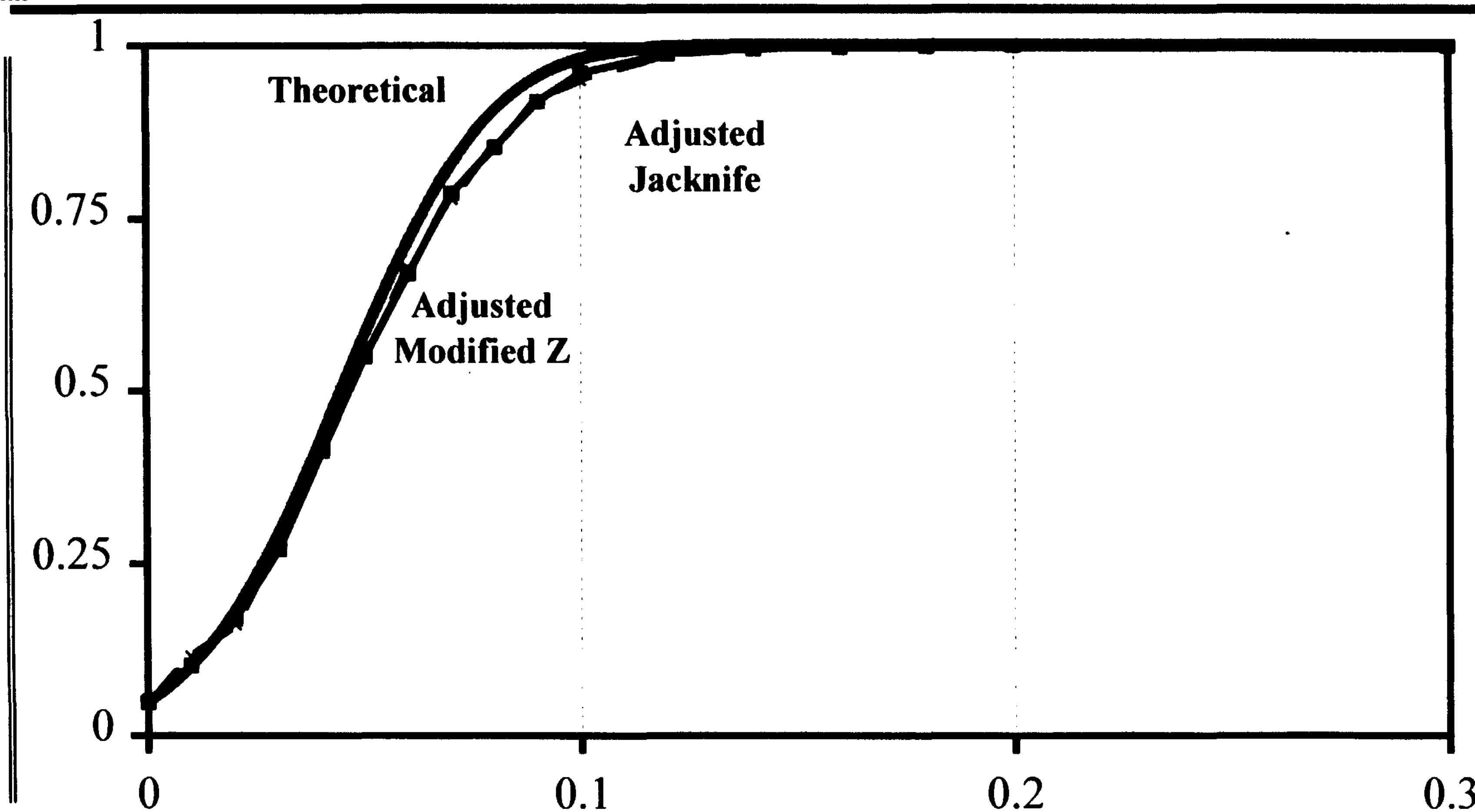
Test	Independent	Medium Correlation	High Correlation
Modified Z	0.559	0.526	0.569
Adjusted Modified Z	0.048	0.149	0.307
Adjusted Jacknife	0.052	0.064	0.059

Total sample size average is 29,120 of which 5% are, on the average, CLEC observations.

Independent Case

1000 Monte Carlo Runs

Critical Value for the Test = -1.65



Total sample size average is 29,120 of which 5% are, on the average, CLEC observations.

econ

Economics Consulting & Quantitative Analysis

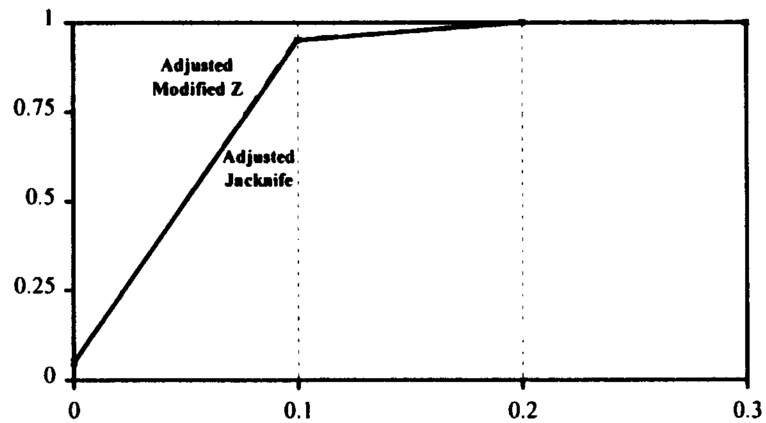
EY/Econ-STAT

100% Violation Power Curves

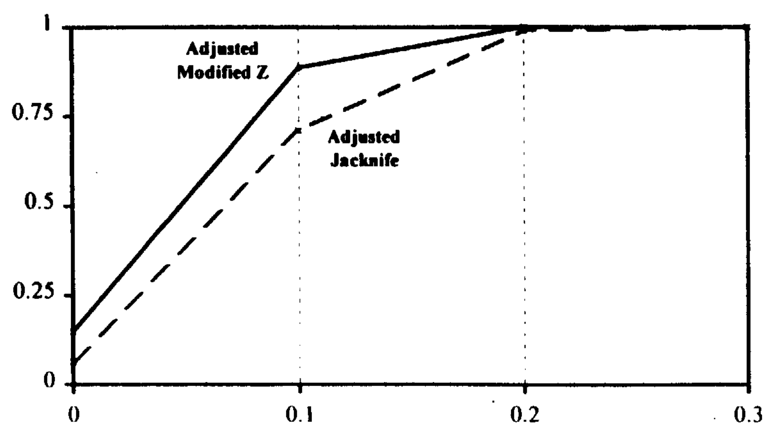
1000 Monte Carlo Runs

Critical value for tests = -1.65

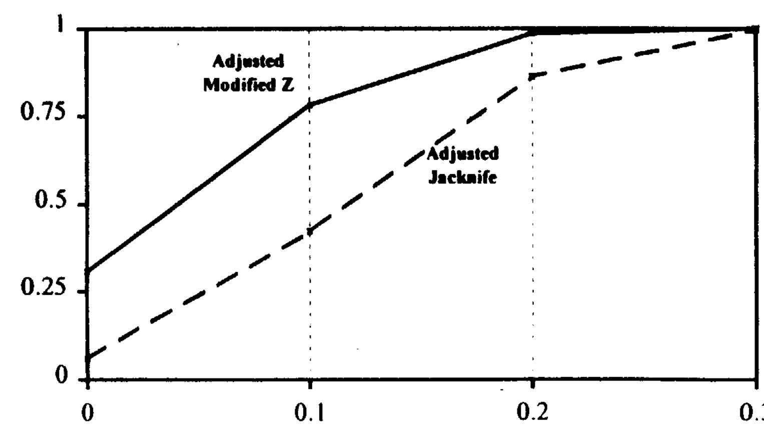
Independent



Medium Correlation



High Correlation



Total sample size average is 29,120 of which 5% are, on the average, CLEC observations.

ERNST & YOUNG LLP

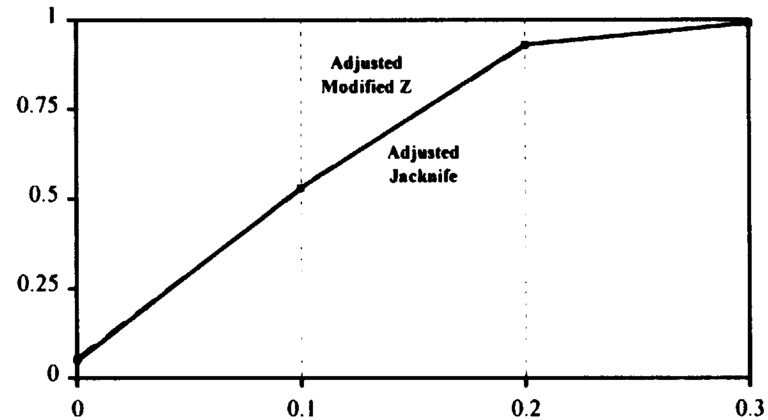
Economics Consulting & Quantitative Analysis
EY/Econ-STAT

50% Violation Power Curves

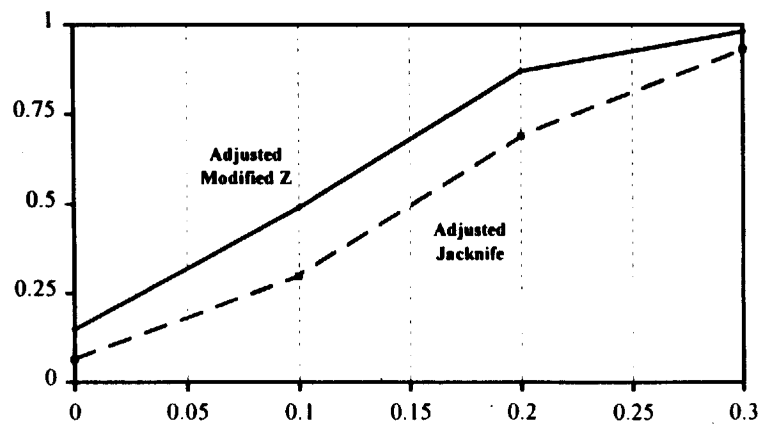
1000 Monte Carlo Runs

Critical value for tests = -1.65

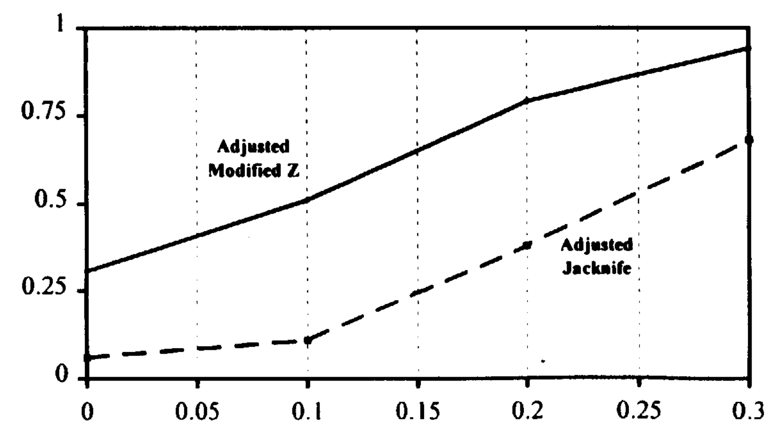
Independent



Medium Correlation



High Correlation



Total sample size average is 29,120 of which 5% are, on the average, CLEC observations.

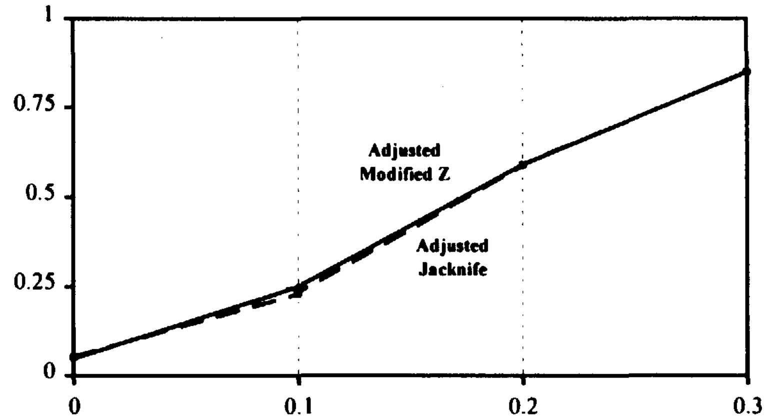
EY econ
Economics Consulting & Quantitative Analysis
EY/Econ-STAT

25% Violation Power Curves

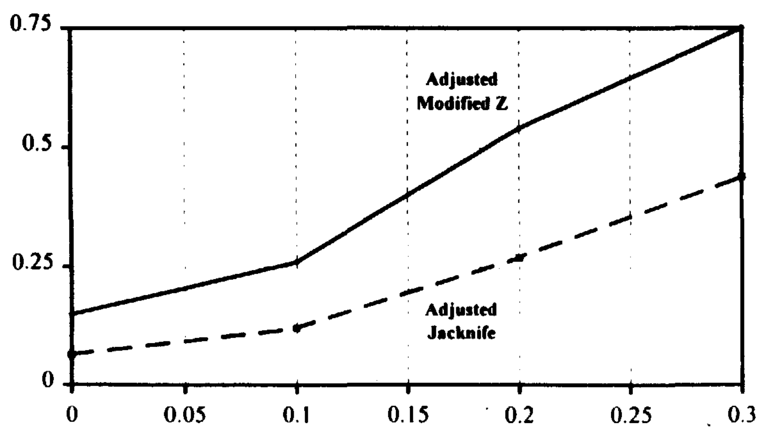
1000 Monte Carlo Runs

Critical value for tests = -1.65

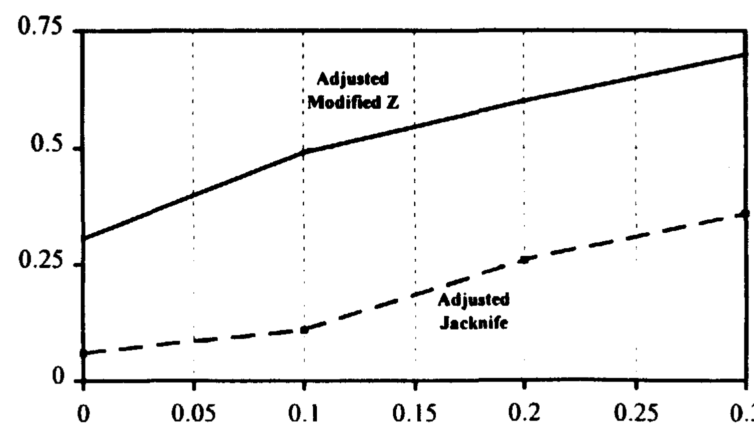
Independent



Medium Correlation



High Correlation



Total sample size average is 29,120 of which 5% are, on the average, CLEC observations.

ERNST & YOUNG LLP

Power of Tests

- Under Independence:
  - Jackknife Test equivalent to Adjusted Modified Z
- Under Correlation:
  - Adjusted Modified Z has wrong Type I rate

Confounding Variables

- Modified Z may be biased when not controlled.
- Controlling reduces bias with little or no loss of efficiency.

Covariance

- BellSouth
  - Jackknife captures total variance.
- LCUG
  - Cannot handle covariance, except by putting burden on ILEC to prove it.

Covariance

- Note that all estimators considered are of the form $\sum_k c_k d_k$
- Components of Variance:

$$Var(\sum_k c_k d_k) = \sum_k c_k^2 Var(d_k) + \sum_k \sum_{j \neq k} c_k c_j Cov(d_k, d_j)$$

where $Cov(d_k,d_j)$ is the covariance between the two quantities.

Components of Variance

- Under a Random Effects Model
  - The first term on the right is within subclass variance.
  - The second is between subclass variance.
- Jackknife method captures between subclass variance within wire centers
- Between subclass variance across wire centers is reduced by treating wire centers as independent when forming replicates.